Disciplina: Descoberta de Conhecimento em Bases de Dados I – Prof. Duncan

Enunciado do trabalho prático com Algoritmos Genéticos

Enunciado do 3º trabalho prático da disciplina, com algoritmos genéticos:

O problema do Fanático por futebol na Copa 2014.

Implementar um algoritmo genético para resolver o problema do Fanático por futebol na Copa 2014.

O fanático por futebol quer ver o máximo de jogos da Copa 2014. Portanto, quer ver 25 jogos! Por simplificação, considera-se que o Fanático não consegue ver 2 partidas no mesmo dia.

De acordo com a tabela da Copa 2014 (material disponível no Moodle), tem-se 25 dias com jogos da Copa (12 a 26/6, 28/6 a 1/7, 4/7, 5/7, 8/7, 9/7, 12/7 e 13/7). Destes, 5 têm apenas 1 jogo previsto:

- 12/06 Brasil e Croácia no Itaquerão
- 8/7 Semifinal 1 no Mineirão
- 9/7 Semifinal 2 no Itaquerão
- 12/7 3º lugar no Mané Garrincha
- 13/7 Final no Maracanã

São 12 estádios em 12 diferentes cidades. Por simplificação, considere que os custos das passagens aéreas são diretamente proporcionais às distâncias aéreas entre as cidades. Ver em <a href="http://www.itatrans.com.br/distancia.html">http://www.itatrans.com.br/distancia.html</a>

Para não dar viés à solução, considerar que o custo da viagem depois do jogo de 12/6 para a 2ª cidade é fixo, bem como o depois do jogo do dia 5/7 para Belo Horizonte (Semifinal 1).

O problema é achar a solução de menor custo de passagens aéreas, assistindo 1 jogo por dia (logo, a função de fitness dá melhores valores quanto menor for a solução global, ou seja, quanto menor for o custo para se deslocar entre as cidades nos 20 dias de jogos não únicos).

Cada cromossomo deve ser codificado com uma lista contendo 20 números entre 1 a 4, onde cada posição da lista representa um dia de jogo. Exemplo: 1ª posição representa 13/6 e pode assumir: 1 – Natal, 2 – Salvador, 3 – Cuiabá; 2ª posição representa 14/6 e pode assumir: 1 – Mineirão, 2 – Recife, 3 – Fortaleza, 4 – Manaus. E assim por diante. Fica a critério de vocês usarem ou números decimais ou binários. Só cuidem no impacto nas operações genéticas.

Gerem uma população inicial de 1000 indivíduos.

Usem roleta viciada para ponderar a qualidade das soluções, e usem elitismo (= 1) entre gerações (o melhor indivíduo de uma geração - melhor fitness - passa para a geração seguinte).

As operações de crossover e mutação são as padrões. Só verifiquem se, para cada dia, qual o número de jogos para codificação das cidades (tem dias com 2, 3 e 4 jogos).

Adotem 80% e 90% para crossover, e 1% e 5% para mutação, e tentem com 1000 gerações.

Analisem o melhor indivíduo encontrado em relação às seleções: Qual ou quais seleções tiveram jogos assistidos pelo Fanático? Quantos foram os jogos por seleção? Qual a seleção mais assistida? E quantos jogos da Seleção Brasileira o Fanático assistiu? Se fosse definir uma nacionalidade para o Fanático, qual seria?

Material a ser entregue: relatório contendo todo o exercício, mais uma reflexão de 30 linhas sobre os achados.

Data de entrega do relatório: 23/06.

No dia 27/06, nos horários propostos, vocês apresentarão o programa funcionando.